Maximization Problems for the Independent Cascade Model

László Hajdu, Miklós Krész, László Tóth

While the research of infection models for networks (directed graphs) is a classic field, its business application has only become widespread in recent years. The definition of the Independent Cascade Model in [1] is the most important breakthrough from a computational theory point of view. The infection spreads the following way: we choose an initial infected vertex set, and only those vertices spread the infection after this that became infected in the previous iteration. Every edge starting from an infected vertex has a single chance for infecting using its own infection probability. Kempe et al. [1] defined the influence-maximization problem, where we are looking for the vertex set for which the expected value of the infected vertices is the highest. This optimization problem is NP-complete, but Kempe et al. proved in [1] that applying a greedy algorithm gives a guaranteed precision that results in good quality solutions in practice. Bóta et al. [2] extended the model and introduced the Generalized Independent Cascade Model. The initial state of this model does not consist of *infected* and *uninfected* vertices, but introduces an *a priori* probability of infection for each vertex. By the end of the infection process, each vertex receives an *a posteriori* infection value. Because the infection process above is #P-complete, [2] introduces a number of approximation algorithms. However, the influence-maximization problem has not been defined for this more general case. In this talk we define two infection-maximization problems connected to the generalized model above. We also present solution methods based on greedy heuristics, and compare their results with each other. We prove that a guaranteed approximation precision can be achieved for these greedy algorithms. To guarantee efficiency from a practical point of view, we decrease the search space for the greedy methods using different approaches (e.g. selection based on vertex degree). We also present a new method that chooses the vertices during the infection process by applying metrics based on the communities (dense subgraphs) of the graph.

References

[1] D. Kempe, J. Kleinberg, E. Tardos Maximizing the Spread of Influence through a SocialNetwork, *Proceeding of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 137-146 (2003)

[2] A. Bóta, M. Krész, A. Pluhár Approximations of the Generalized Cascade Model, *Acta Cybern. 21*, 37-51 (2013)